HAL
open science

Invariant methods for an ensemble-based sensitivity analysis of a passive containment cooling system of an AP1000 nuclear power plant

Francesco Di Maio, Giancarlo Nicola, Emanuele Borgonovo, Enrico Zio

To cite this version:

Francesco Di Maio, Giancarlo Nicola, Emanuele Borgonovo, Enrico Zio. Invariant methods for an ensemble-based sensitivity analysis of a passive containment cooling system of an AP1000 nuclear power plant. Reliability Engineering and System Safety, 2016, 151, pp.12-19. 10.1016/j.ress.2015.10.006 . hal-01786978

HAL Id: hal-01786978

https://hal.science/hal-01786978v1

Submitted on 23 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Invariant Methods for an ensemble-based Sensitivity Analysis of a Passive Containment Cooling System of an AP1000 Nuclear Power Plant

Francesco Di Maio[1*], Giancarlo Nicola[1], Emanuele Borgonovo[2], Enrico Zio[1,3]

[1]*Energy Department, Politecnico di Milano, Via La Masa 34, 20156 Milano, Italy*

francesco.dimaio@polimi.it

[2]*Department of Decision Sciences Bocconi University, Milano, Italy*

[3]*Chair on Systems Science and the Energetic Challenge, Fondation EDF, Ecole Centrale Paris and Supelec, France*

ABSTRACT: Sensitivity Analysis (SA) is performed to gain fundamental insights on a system behavior that is usually reproduced by a model and to identify the most relevant input variables whose variations affect the system model functional response. For the reliability analysis of passive safety systems of Nuclear Power Plants (NPPs), models are Best Estimate (BE) Thermal Hydraulic (TH) codes, that predict the system functional response in normal and accidental conditions and, in this paper, an ensemble of three alternative invariant SA methods is innovatively set up for a SA on the TH code input variables. The ensemble aggregates the input variables raking orders provided by Pearson correlation ratio, Delta method and Beta method. The capability of the ensemble is shown on a BE-TH code of the Passive Containment Cooling System (PCCS) of an Advanced Pressurized water reactor AP1000, during a Loss Of Coolant Accident (LOCA), whose output probability density function (pdf) is approximated by a Finite Mixture Model (FMM), on the basis of a limited number of simulations.

KEYWORDS: Invariant sensitivity analysis methods; Pearson correlation ratio, Delta method; Beta Method; AP1000 Nuclear Power Plant; Ensemble of methods.

1 INTRODUCTION

The reliability analysis of passive safety systems of advanced Nuclear Power Plants (NPPs) must consider that when uncertainties of counter-forces (e.g., friction) have magnitude comparable to the driving ones (e.g., gravity, natural circulation), physical phenomena may fail performing the intended function even if i) safety margins are met, ii) no hardware failures occur (Marques et al. 2005, Burgazzi 2007, Zio et al. 2009).

Many approaches have been proposed for identifying and quantifying the uncertainties affecting the code outputs and generated by simplifications, approximations, round-off errors, numerical techniques, user errors and variability in the input parameters values (Pourgol-Mohammad 2009) e.g., Code Scaling, Applicability, and Uncertainty (CSAU) (Wulf et al. 1990),

Automated Statistical Treatment of Uncertainty Method (ASTRUM) and Integrated Methodology for Thermal-Hydraulics Uncertainty Analysis (IMTHUA) (Glaeser et al. 1994). All these methods deal with the need of addressing the problem of uncertainty quantification of the Thermal Hydraulics (TH) codes that are used to predict the response of the systems in nominal and accidental conditions. Traditionally, these calculations were performed on very conservative TH codes, that were supposed to "cover" the system from undesired and/or unknown system (uncertain) behaviors (Zio et al., 2008). More recently, Best Estimate (BE) codes have been adopted to provide more realistic results, thus avoiding over-conservatism (Zio et al. 2010, 10CFR50.46), although, requiring a detailed, precise and rigorous treatment of the related uncertainties.

This has brought into the reliability analysis of NPPs passive systems an increasing computational complexity that has been recently addressed in literature: for example, a combination of Order Statistics (OS) (Guba et al. 2003, Zio et al. 2008a) and Artificial Neural Networks (ANN) has been proposed to speed up the computations (Secchi et al. 2008). However, these approaches allow determining only some percentiles and not the whole distribution, and do not provide insights on the sensitivity to input variability (Langewisch 2010; Hong et al. 2011).

In this respect, several SA methods have been proposed (Saltelli et al., 2002): some are quantitative and model-free, whereas some others are specifically tailored to the model. Among those belonging to the former group, global SA methods offer great capabilities but, again, high computational costs, especially if compared to local and regional SA methods. The most used global SA methods are: non parametric methods, variance-based methods, moment independent, value of information-based methods and Monte carlo filtering (the interested reader may refer to (Borgonovo et al., 2015) for a detailed review of the methods). Examples of global non parametric SA methods are the Standardized Regression Coefficient (Helton, 1993) and the Pearson coefficient (Saltelli et al., 1990). The functional ANOVA expansion of the input-output mapping (Helton 1993, Saltelli et al. 2000, Cadini et al. 2007) is at the basis of variance-based methods, which are widely used in global SA. However, ANOVA expansion requires independence among the model inputs and, if the number of parameters is high, a high computational cost is required for computing interaction terms (Yu et al., 2010). Variance-based sensitivity measures have been originally defined by Pearson in 1905 and are known under the name of correlation ratio and have been further improved by Sobol in 1993. In general terms, when used as measures of statistical dependence, first order variance-based sensitivity measures as well as

non-parametric methods may lead to misleading conclusions, especially when model inputs are correlated. These limitations are overcome by moment independent sensitivity measures. Among these, the invariant method Delta (Borgonovo 2007) and Kolmogorov-Smirnov distance between cumulative distribution functions (Baucells and Borgonovo, 2013) have to be cited as viable solutions.

In this work, to avoid a large number of TH code runs for the numerical estimation of the selected sensitivity measures, we propose an innovative framework of analysis whose flowchart is shown in Figure 1. The idea is to directly rely on the information available in the multimodal pdf of the output variable for performing global SA of a TH code. First, a limited number N of simulations of the TH code are performed and a Finite Mixture Model (FMM) is used to reconstruct the pdf of the model output (Carlos et al., 2013). The natural clustering made by the FMM on the TH code output (McLachlan et al. 2000, Di Maio et al. 2014a) is exploited to estimate global sensitivity measures using a given data approach (Plischke et al, 2013). As shown in (Borgonovo et al (2014)), in fact, variance-based and distribution-based sensitivity measures rest on a common rationale that allows them to be estimated from the same design of experiments. We can, then, employ an ensemble of three SA indicators: first-order variance-based sensitivity measure (i.e., the Pearson's correlation ratio), the Delta method (Borgonovo 2007) and a new sensitivity measure based on the Kolmogorov-Smirnov distance between cumulative distribution functions (Baucells and Borgonovo, 2013). We use these sensitivity measures for ranking the input variables most affecting the output uncertainty. The ensemble strategy allows combining the output of the three individual methods (that perform more or less well depending on the data) to generate reliable rankings (Di Maio et al. 2014b). The idea of using an ensemble of methods for sensitivity analysis will be shown particularly useful when the number of TH code simulations is small, for a low computational cost: due to the limited quantity of data in this situation, in fact, possible misleading rankings can arise from the individual SA methods, whereas the diversity of the methods integrated in the ensemble allows overcoming the problem.

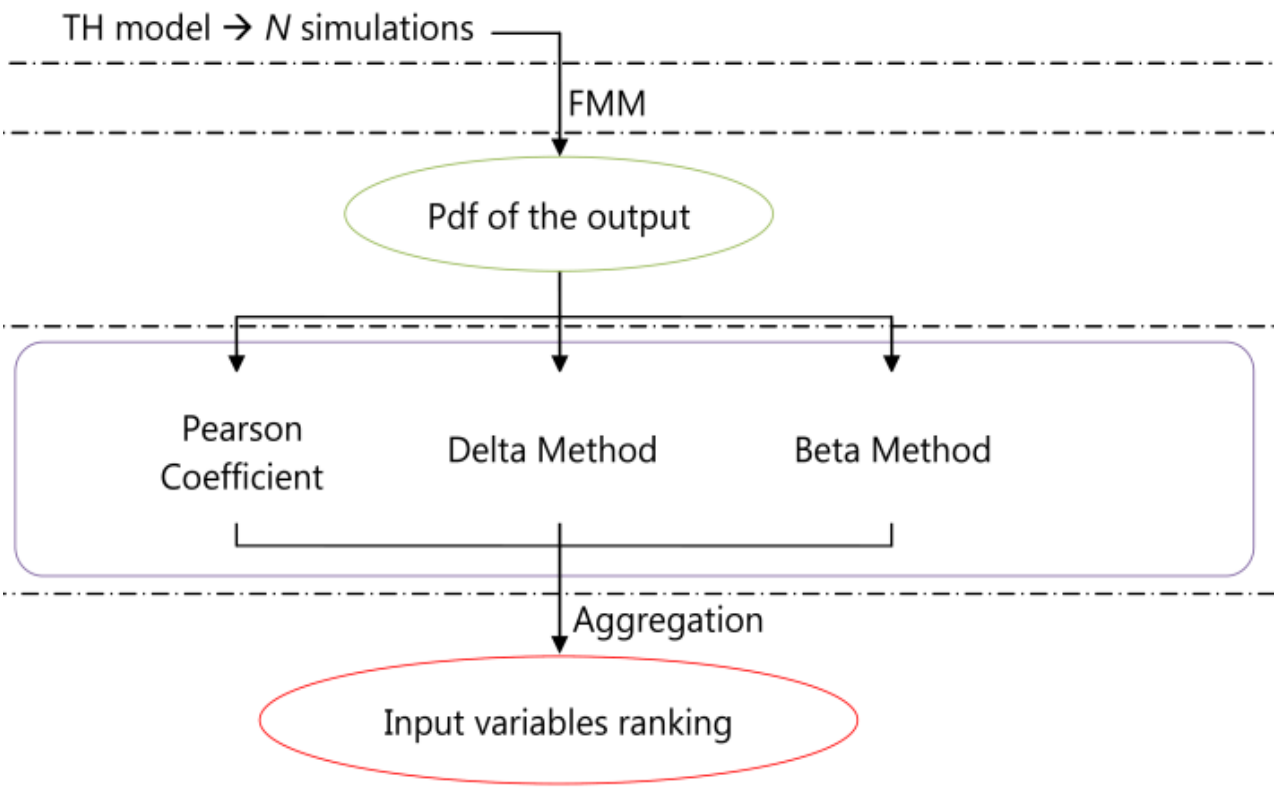

Figure 1. Flowchart of the proposed framework for ensemble Sensitivity Analysis.

Our application concerns the sensitivity analysis of a TH code that simulates the behaviour of the Passive Containment Cooling System (PCCS) of an Advanced Pressurized water reactor AP1000 during a Loss Of Coolant Accident (LOCA). The combination of the three sensitivity methods is shown to make the results robust, with no additional computational costs (no more TH code runs are required for SA).

The paper is organized as follows. In Section 2, the case study and the relative TH code are illustrated. In Section 3, the basis of FMM are presented along with the ensemble of sensitivity methods, i.e., Pearson correlation, Delta method and Beta method. In Section 4, the experimental results are reported. Section 5 draws some conclusions.

2 CASE STUDY

The AP1000 NPP is a 1117 MWe (3415 MWth) pressurized water reactor (PWR), with a passive Residual Heat Removal System (RHRS) and a Passive Containment Cooling System (PCCS). The PCCS cools the containment following an accident, so that pressure is effectively controlled within the safety limit of 0.4 MPa. During an accident (for example, during a Loss

Of Coolant Accident (LOCA) or a Main Steam Line Break (MSLB) accident), the produced steam is injected into the containment and *i*) an air baffle incorporated into the concrete structure outside the steel vessel creates the tunnel for continuous, natural circulation of air, and *ii*) water that drains by gravity from a tank located on top of the containment shield building (by means of three redundant and diverse water drain valves) supplements, by evaporation, the heat removal. The steel containment vessel provides the heat transfer surface through which heat is removed from inside the containment and transferred to the cold sink, i.e., the atmosphere. In addition, even in case of failure of water drain, air-only cooling is supposed to be capable of maintaining the containment below the failure pressure (Schulz 2006). Figure 2 shows the PCCS of the AP1000 (Westinghouse Electric Company).

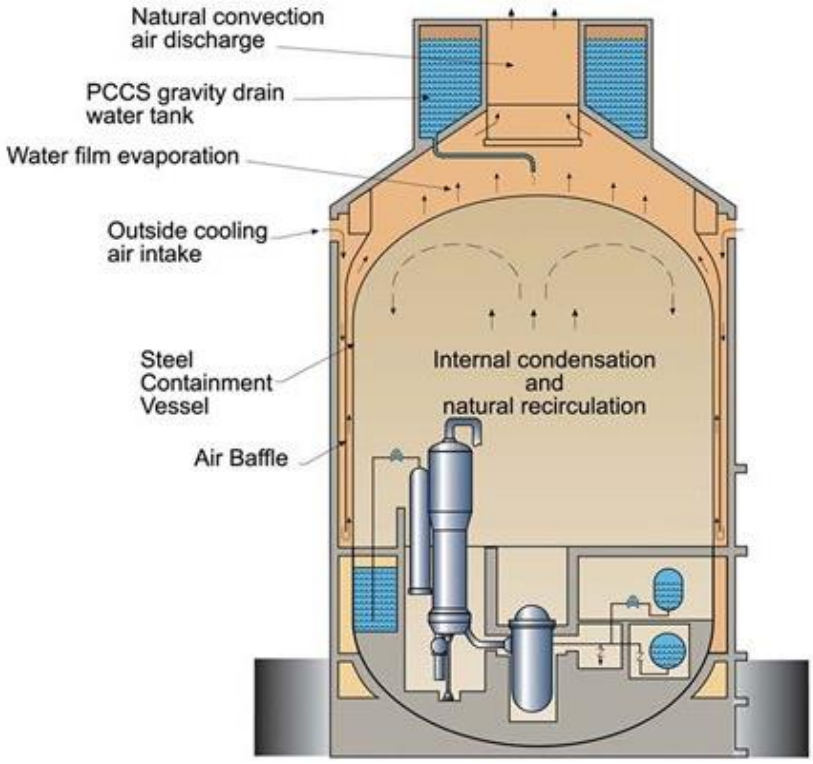

Figure 2. AP1000 Passive Containment Cooling System (Westinghouse Electric Company).

For the quantification of the functional failure of the PCCS of the AP1000 following a LOCA, a TH model for stratified heat transfer with non-condensed gas has been developed, that consists in four phases (Rahim et al. 2012): 1) blowdown, from the accident initiation (by a double-ended guillotine pipe break in a primary coolant line affecting the normal operation of the reactor at steady-state full power) to the time at which the primary circuit pressure reaches the containment pressure; 2) refill, from the end of the blowdown to the time when the Emergency Core Cooling System (ECCS) refills the vessel lower plenum; 3) reflood, which begins when

water starts flooding the core and ends when this is completely quenched; 4) post-reflood, which starts after the core quenching and during which energy is released to the Reactor Coolant System (RCS).

In the post-reflood phase, the steam produced in the RCS is cooled at the internal face of the steel containment vessel and, then, the heat is conducted by the vessel and transferred to the atmosphere in the air channels (see Figure 2). Cold air enters the channels through the three rows of air inlets and flows down to the bottom of the channels, where it is heated by the steel vessel up to the air diffuser to the environment.

The TH code simulates the dynamics of the heat transfer processes in the post-reflood phase with a stratified dome of the PCCS vessel. The output variable is the pressure value of $P_{containment}$ after 1000 s from the initiation of the LOCA. The D=51 input variables are listed in the Appendix A together with their distributions chosen from expert judgment and literature review (Zio et al. 2008b, Zio et al. 2010; Yu et al., 2015). Three families of distributions have been used: seasonal, normal and uniform. Seasonal relates to the external air temperature T_{inlet} and pressure P_{inlet} variability, as inferred by historical data collected by a representative Chinese Automatic Weather Station (CAWS) in different months. Normal distributions, e.g., for the LOCA steam temperature, T_{steam}, are truncated distributions with mean μ and support equal to 4σ, where σ is the standard deviation. For uniform distributions, e.g. for the friction factors, the supports from "Lower value" to "Upper value" are reported.

It is worth mentioning that most of the considered inputs can be considered as independent of each other, but some correlation among a few of them should be considered and handled for a proper identification of the most relevant input variables in affecting the containment pressure $P_{containment}$. The initial air temperature T_{inlet} and pressure P_{inlet} in the containment are mutually influenced by the local climate where the NPP is built by the operational conditions. For example, NPPs located in colder places will be favored in terms of reliability of the PCCS because of the benefit that can be gained by T_{inlet} lower than if located in hotter places; moreover, under normal operational conditions, the within-containment temperature is slightly larger than T_{inlet} and the within-containment pressure is slightly smaller than P_{inlet}. These considerations (i.e., a mixture of input variables that are partially correlated (Yu et al., 2015) and independent) has suggested the idea of resorting to an ensemble of global SA methods that would be able to deal with both hypotheses (either correlated or independent inputs), where each one of the 3 implemented methods is best suited for one of the two hypotheses.

3 UNCERTAINTY AND ENSEMBLE-BASED SENSITIVITY ANALYSES

Let y denote the output of a TH model m, viz:

$$y = m(x_1, x_2, \dots, x_l, \dots, x_D) \quad l = 1, \dots, D \quad (1)$$

where x_l is the l-th input variable. The random output variable y follows a finite mixture density $f(y)$ with K models if (Carlos et al., 2013, Di Maio et al. 2014b):

$$f(y) = \sum_{k=1}^{K} \pi_k f_k(y|\theta_k) \quad (2)$$

where $f_k(y|\theta_k)$ are K different probability density functions, θ_k is the set of parameters of the k-th model of the mixture and π_k are the mixing probabilities that satisfy:

$$\sum_k \pi_k = 1 \qquad with\ \forall\, k, \pi_k \geq 0 \quad (3)$$

In particular, if $f_k(y|\theta_k)$ is Gaussian, then:

$$f_k(y|\theta_k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(y-\mu_k)^2}{2\sigma_k^2}} \quad (4)$$

where $\theta_k = (\mu_k, \sigma_k)$ are the mean and the standard deviation of the k-th gaussian mixture model, respectively.

The parameters θ and π of the mixture models have to be found by Expectation Maximization (EM) aimed at best approximating the pdf of the output y of the TH model with a small number N of TH code simulations. In addition, "natural" clusters corresponding to each Gaussian model $f_k(y|\theta_k)$ of the mixture are found: some may be representative of normal conditions whereas others of accidental conditions, allowing for a direct calculation of the probability of exceeding a certain safety limit (i.e., of system functional failure). These clusters will be exploited for SA within an ensemble of the three methods of Pearson' correlation ratio, Delta method and Beta method, whose individual outcomes are aggregated for identifying the input variables most affecting the output uncertainty (as in Figure 1).

For completeness, we summarize below the basics of sensitivity methods considered for the ensemble. Let Y be the output of interest and X_i a generic model input. Pearson's correlation ratio (Plischke et al, 2013) is the first-order variance-based sensitivity index,

$$\eta_i = \frac{V[E\{Y \mid X_i\}]}{V[Y]} \qquad (5)$$

η_i was introduced by Pearson as a stronger measure of statistical dependence than the classical statistical correlation coefficient. In the case model inputs are independent, η_i coincides with Sobol' index of order 1. However, it has been shown that a model output Y can depend on a model input X_i although the value registered by η_i is null. That is, we are not reassured that a null value of η_i indicates that the model output is independent of X_i. Then, one can use stronger measures of dependence looking at the distance between the distribution of the model output and the conditional distribution obtained once X_i is fixed at any of its values. Here, we consider two representatives of the class of distance-based sensitivity measures. The first is the δ-importance measure (Borgonovo 2007):

$$\delta_i = \frac{1}{2} E[d_i(X_i)] \qquad (6)$$

where

$$d_i(x_i) = \int_{\Omega_Y} \left| f_Y(y) - f_{Y|X_i}(y) \right| dy \qquad (7)$$

is the separation (i.e., the area enclosed) between the density of Y, $f_Y(y)$ and the density conditional on given X_i, $f_{Y|X_i}(y)$. The δ-importance measure possesses several convenient properties. It is monotonic transformation invariant and it does not require independence between the model inputs. Indeed, in (Plischke et al., 2013), it is shown that δ_i is a stronger measure of statistical dependence between X_i and Y than η_i: it can be proven that $\delta_i = 0$ if and only if Y is independent of X_i, that is, if δ_i=0, the analyst is reassured that the model output Y is independent of X_i. This latter condition, as already discussed, does not hold for η_i. We also use the Beta measure [Baucells and Borgonovo (2013)]:

$$\beta_i = E[\sup_{y \in \Omega_Y} \left| F_Y(y) - F_{Y|X_i}(y) \right|] \qquad (8)$$

which is a transformation-invariant sensitivity measure based on the Kolmogorov-Smirnov distance between the cumulative distribution functions of Y (here denoted with $F_Y(y), F_{Y|X_i}(y)$).

The intuition that η_i, δ_i and β_i can be written, in general terms, as the sensitivity index ξ_i [Borgonovo et al 2014b]

$$\xi_i = E\left[\gamma_i\left(X_i\right)\right] \qquad (8)$$

where $E\left[\gamma_i\left(X_i\right)\right]$ is an outer expectation over the marginal distribution of the model input X_i and γ_i is an inner statistic

$$\gamma_i\left(x_i\right) = \varsigma\left(F_Y, F_Y \middle| X_i = x_i\right) \qquad (9)$$

clarifies the reason why these sensitivity measures can all be estimation from the same design. The interested reader, may refer to Baucells and Borgonovo (2013), Borgonovo et al (2014b) and (Borgonovo et al., 2015) for a more detailed discussion on the properties of η_i, δ_i and β_i. In (Borgonovo et al, 2014a), it is shown that given data estimators based on the above-mentioned common rationale are consistent estimators of any global sensitivity measure. In this work, the ensemble of sensitivity methods is embedded within a FMM, for speeding up the computation of the measures η_i, δ_i and β_i for a limited number N of BE-TH codes simulations.

3.1 *Ensemble-based SA*

For the evaluation of the sensitivity of the output variable on the input variables, the outputs of the three SA methods considered are aggregated in an ensemble output, without requiring any additional TH simulations. The ensemble outcome is expected to provide more accurate and confident rankings. Majority voting R_{mv} (Di Maio et al. 2012) and R_{sum} aggregation (Kukkonen et al. 2007) have been considered. The former consists in taking the majority voting of the SA ranks provided by the three methods: the ranking orders (each one of length D=51) provided individually by the three methods are aggregated by assigning to each ranking position $R_{mv,l}$ the l-th input voted by majority; if for one position none of the methods agree, no input is ranked in that ranking position. The main advantages of this aggregation are its simplicity and negligible computational burden. The R_{sum} aggregation (Kukkonen et al. 2007) consists in taking for each l-th input variable the sum of the ranking positions provided by the individual methods and then sorting them with respect to $R_{sum,l}$.

4 RESULTS

The SA ensemble approach described in Section 3 has been applied to the case study of Section 2. In order to test the effect of different amounts of available information, sample sizes of LOCAs are simulated with N=1000, 600, 400, 200, 100 samples of the input variables values drawn from the distributions reported in Appendix A.

As an example, in Figure 3 the distribution (histogram) of the output variable $P_{containment}$ and its FMM (line) of two Gaussian distributions $f_k(\theta_k),\ k = 1,2,$ is shown for N=1000 and 100 simulations. The parameters of the mixture models found for different sizes of N=1000, 600, 400, 200, and 100 are reported Table 1.

The choice of using two Gaussian distributions for the FMM parameters identification is the result of a trial and error procedure. It can be shown that the multinomial pdf of $P_{containment}$, $f(P_{containment})$, is well reconstructed by the FMMs built with a number of code runs from N=1000 down to N=100 (Di Maio et al, 2014a,b).

Table 1 Parameters of the finite mixture of Gaussian distributions computed with the EM algorithm for different sample sizes

N	Probabilities (π_1, π_2)	Means (μ_1, μ_2)	Standard deviations (σ_1, σ_2)
1000	(0.93, 0.07)	(0.3715, 0.4006)	(0.0119, 0.0047)
600	(0.93, 0.07)	(0.3720, 0.4015)	(0.0121, 0.0052)
400	(0.92, 0.08)	(0.3727, 0.4022)	(0.0119, 0.0057)
200	(0.94, 0.06)	(0.3707, 0.3993)	(0.0122, 0.0025)
100	(0.93, 0.07)	(0.3717, 0.3998)	(0.0119, 0.0030)

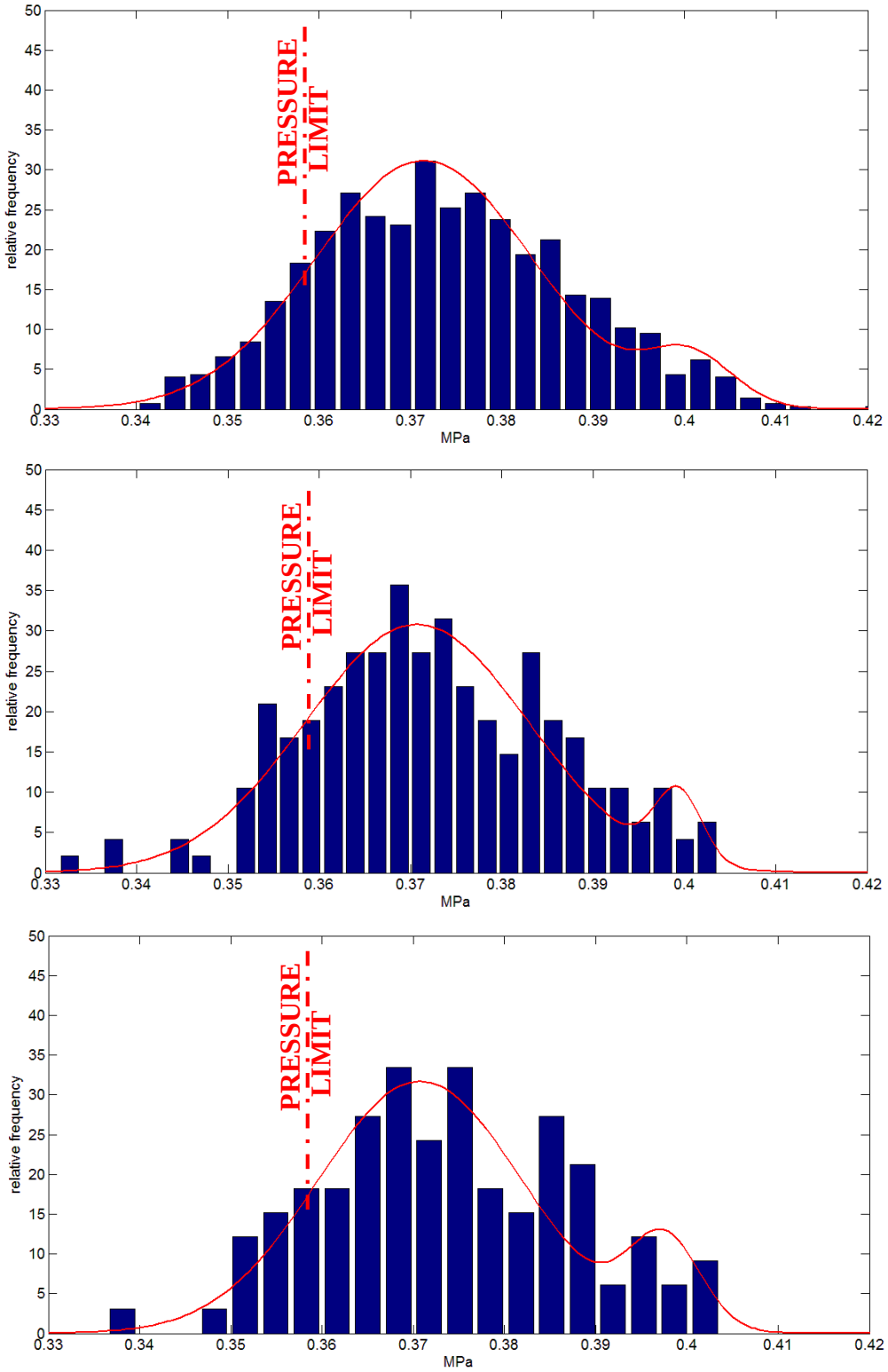

Figure 3 Histogram of the model output and mixture model reconstruction with N=1000 code runs (top), N=200 (middle) and with N=100 (bottom).

4.1 *Sensitivity Analysis*

The three SA methods illustrated in Section 3 have been applied exploiting the analytical pdf obtained by the FMMs and the related clustering. In general terms, the contribution of the input variables in shaping the K clusters in which the analytical output pdf, obtained by the

FMM, can be represented is related to the rankings provided by each SA method. Then, their ensembles are analyzed in the following, for the different cases of code runs from N=1000 to N=100.

The ranking values provided by the Pearson's correlation ratio, and Delta and Beta methods can be shown to be robust when applied to N=1000, 600, 400 and 200 code runs of the lumped-parameters (i.e., simplified and fast-running) TH model that is used to analyze the effect of the inputs on the output reactor power on the PCCS function (Di Maio et al. 2014c). In other words, for this simplified TH model design, whose computational demand is significantly smaller than for the TH model of Section 2, the two most important parameters are found to be T_{inlet} and G independently from the number N and even for a small sample size of N=200. Therefore, for the case study of Section 2, we adopt the number N=200 as limiting condition for the individual application of the three SA methods. Table 2 shows the ranking obtained (with both aggregation R_{mv} and R_{sum}), when the number of simulated LOCAs with the complex TH model of Section 2 is equal to N=200 (Table 2).

Table 2 Input variables ranking obtained with N=200 TH code runs

	1°	**2°**	**3°**	**4°**
Pearson	T_{inlet}	G	α_1	D_3
Delta	T_{inlet}	D_2	Cp_2	G
Beta	T_{inlet}	α_1	G	D_3
Ensemble R_{mv}	$\boldsymbol{T_{inlet}}$	$\boldsymbol{G/D_2/\alpha_1}$	$\boldsymbol{\alpha_1/Cp_2/G}$	$\boldsymbol{D_3}$
Ensemble R_{sum}	$\boldsymbol{T_{inlet}}$	$\boldsymbol{G}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{D_3}$

In general terms, the most important variables are found to be T_{inlet}, G, α_1 and D_3. The input variables identified as important play a direct role in the heat transfer process of the PCCS operation: the temperature of the ultimate heat sink T_{inlet} is related to the capability of the atmosphere to absorb the heat generated, the steam mass flow rate G is related with the energy entering the containment, the conductivity of the containment α_1 has an active role in determining the heat flux leaving the containment, and the containment diameter D_3 determines the internal volume and the heat exchange surface, both critical quantities for internal pressure and heat transfer.

Input variable T_{inlet} is identified as the most important one by all the three alternative methods. The Beta method switches the positions of G, α_1 and D_3 with respect to Pearson's correlation ratio, while Delta suggests D_2 and Cp_2 as important inputs.

The ensemble of methods R_{sum} overcomes the problem of disagreement among the three methods providing the order: T_{inlet}, G, α_1, D_3. On the other hand, R_{mv} is not able to univocally assign the proper ranking to the variables. For example, G, D_2 and α_1 are all scored as second most important variables, because none of these gains more than one vote among the Pearson's correlation ratio, Delta and Beta methods. Despite that, the comparison of the two rankings obtained with R_{mv} and R_{sum} provides the analyst with more robust information: the reliance on only one of these methods might induce the analyst to mistrust the ranking outcome due to the shortage of TH code runs (i.e., when using the Pearson's correlation ratio in cases of weak statistical correlation between the inputs and the output). This is in line with best practices, that recommend that analyst do not rely on a sole sensitivity measure, especially in the presence of a sample of limited size.

As a comparison, it is worth noticing that ranking order T_{inlet}; G, α_1 and D_3 is the same that can be obtained with N=1000 within a FMM-based SA (Di Maio et al. 2014a) and with N=1000 with both aggregation R_{mv} and R_{sum}. This result shows that the aggregation of the Pearson's correlation ratio, and Delta and Beta methods can leverage the computational burden from N=1000 (~275 hours on an Intel Core2Duo P7550) to N=200 (~55 hours)

As a final remark and for ease of clarity, let us now suppose that only N=100 TH code runs are available. Table 3 shows the results obtained with N=100 TH code runs. The ensemble based on R_{mv} still identifies the previous four variables as most important but cannot produce the same ranking order. This is due to the the very limited sample size, that induces instability in the estimates of the sensitivity measures. Suppose now that an analyst was forced at this small sample: how to proceed? A first way would be also to use R_{sum}. A second way is to use the bootstrap method for obtaining a rubust ranking estimates and, thus, being able to order the input variables with a degree of confidence in the results (Di Maio et al., 2015).

As for some concluding remarks drawn from the case study analyzed, it appears that: i) the ensembles proposed are effective with small numbers N (e.g. N=200), compensating the errors of the individual methods, ii) the R_{sum} aggregation strategy has provided the best results, especially with N=200.

Table 3 Input variables ranking obtained with N=100 TH code runs

	1°	2°	3°	4°
Pearson	T_{inlet}	D_3	G	α_1
Delta	T_{inlet}	A_2	α_1	f_8
Beta	T_{inlet}	D_3	G	α_1
Ensemble R_{mv}	$\boldsymbol{T_{inlet}}$	$\boldsymbol{D_3}$	$\boldsymbol{G}$	$\boldsymbol{\alpha_1}$
Ensemble R_{sum}	$\boldsymbol{T_{inlet}}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{D_2}$	$\boldsymbol{G}$

5 CONCLUSIONS

In this paper, we have proposed a SA framework with low computational burden. The framework here proposed consists in i) estimating a Gaussian FMM to retrieve the analytical pdf of the model output with as few simulations as possible ii) induce a clustering of the output variable space iii) applying two (alternative) ensemble strategies to aggregate three sensitivity methods, namely Pearson's correlation ratio, Delta method and Beta method.

The framework has been tested on a TH code that simulates the behaviour of a Passive Containment Cooling System (PCCS) of an Advanced Pressurized reactor AP1000 during a Loss Of Coolant Accident (LOCA).

The results obtained show the capability of the framework in discerning between influent and negligible input variables at a reasonable computational cost.

Acknowledgements

This research has been partially carried out within the collaboration with the School of Nuclear Science and Engineering of the North China Electric Power University (Beijing, China) that is supported by "The National Natural Science Foundation of China" (51206042).

References

10CFR50.46. Acceptance criteria for emergency core cooling systems for light-water nuclear power reactors, NRC Regulations.

Baucells M. and Borgonovo E., 2013: "Invariant Probabilistic Sensitivity Analysis," *Management Science*, 59 (11), pp. 2536-2549

Borgonovo E, 2006. Measuring uncertainty importance: investigation and comparison of alternative approaches, Risk Analysis, 26, 5, 1349-1362.

Borgonovo E., 2007: "A New Uncertainty Importance Measure", *Reliability Engineering and System Saftey*, 92, pp. 771-784

Borgonovo E., 2013, Sensitivity analysis in decision making, Encyclopedia of operations research and management science, 1-12.

Borgonovo E., Hazen G.B. and Plischke, E., 2014a: "A common Rationale for Global Sensitivity Measures and Their Estimation", MIMEO, submitted.

Borgonovo E., Tarantola S., Plischke E., and Morris M.D., 2014b: "Transformations and Invariance in the Sensitivity Analysis of Computer Experiments," *Journal of the Royal Statistical Society, Series B*, forthcoming

Borgonovo E., Plischke E, 2015. Sensitivity Methods for operational research: a review of recent advances, submitted.

Cadini, F., Zio, E., Di Maio, F., Kopustinkas, V., Urbonas, R. 2007. A Neural-network-based Variance Decomposition Sensitivity Analysis, International Journal of Nuclear Knowledge Management, Vol. 2, No. 3, 2007, pp. 299-312.

Carlos, S., Sánchez, A., Ginestar, D., Martorell, S. 2013. Using finite mixture models in thermal-hydraulics system code uncertainty analysis. Nuclear Engineering and Design, Volume 262, Pages 306-318.

Di Maio, F., Hu, J., Tse, P.W., Pecht, M, Tsui, K.L., Zio, E., 2012. Ensemble-approaches for clustering health status of oil sand pumps, Expert Systems with applications, Vol. 39, pp 4847-4859.

Di Maio, F., Nicola, G., Zio, E., Yu, Y. 2014a. Finite Mixture Models for sensitivity analysis of Thermal Hydraulic Codes for Passive Safety Systems safety Analysis, Nuclear Engineering and Design.

Di Maio, F., Nicola, G., Zio, E., Yu, Y. 2014b. Ensemble-based sensitivity analysis of a best estimate thermal hydraulic model of a Passive Containment Cooling System of an AP1000 Nuclear Power Plant, Annals of Nuclear Energy, 73, 200–210, 2014.

Di Maio, F., Nicola, G., Zio, E., Yu, Y. 2014c. Failure Damage Domain Identification of the Passive Containment Cooling System of an AP1000 Nuclear Reactor, PSAM12, Probabilistic Safety Assessment & Management, 22-27 June 2014, Honululu, USA.

Di Maio, F., A. Bandini, E. Zio, S. Carlos Alberola. 2015. Bootstrapped Ensemble-based Sensitivity analysis of a TRACE thermal-hydraulic model based on a limited number of PWR large Break LOCA simulations, under review.

Glaeser, H., Hofer, E., Kloos, M., Skorek, T. 1994. Uncertainty and sensitivity analysis of a post-experiment calculation in thermal hydraulics. Reliab. Eng. Syst. Safe.45 (1/2), 19–33.

Guba, A., Makai, M., Pal, L., 2003. Statistical aspects of best estimate method-I. Reliab. Eng. Syst. Safe. 80, 217–232.

Helton, J. C., 1993. Uncertainty and Sensitivity analysis Techniques for Use in Performance Assessment for Radioactive Waste Disposal, Reliability Engineering and System Safety, 42, pp. 327-367.

Hong, I.S., Oh, D.Y., Kim, I.G., 2011. Generic application of Wilks tolerance limit evaluation approach to nuclear safety. In: Proceedings of the OCDE/CSNI Workshop on Best Estimate Methods and Uncertainty Evaluations, Barcelona, Spain.

Kukkonen, S. & Lampinen, J., 2007. Ranking-Dominance and Many-Objective Optimization, 2007 Evolutionary Computation, 25-28 Sept., Singapore, pp. 3983–3990.

Langewisch, D.R., 2010. Uncertainty and Sensitivity Analysis for Long-running Computer Codes: A Critical Review. Massachusetts Institute of Technology, http://hdl.handle.net/1721.1/58285

Marques, M., Pignatel, J.F., Saignes, P., D'Auria, F., Burgazzi, L., Muller, C., Bolado-Lavin, R., Kirchsteiger, C., La Lumia, V., Ivanov, I., 2005. Methodology for the reliability evaluation of a passive system and its integration into a probabilistic safety assessment. Nuclear Engineering and Design 235, 2612–2631.

McLachlan, G., Peel, D., 2000. Finite Mixture Models. John Wiley & Sons Inc., New York.

Plischke E., Borgonovo E. and Smith C.L., 2013: "Global Sensitivity Measures from Given Data," *European Journal of Operational Research*, 226:536–550.

Pourgol-Mohammad, M., 2009. Thermal-hydraulics system codes uncertainty assessment: a review of the methodologies. Ann. Nucl. Energy 36, 1774–1786.

Rahim, F.C., Rahgoshay, M., Mousavian, S.K., 2012. "A study of large break LOCA in the AP1000 reactor containment", Progress in Nuclear Energy, Volume 54, Issue 1, Pages 132–137.

Saltelli A., Marivoet, J., 1990. Non-parametric statistics in sensitivity analysis for model output: a comparison od selected techniques, Reliability Engineering and System Safety, 28, 2, 229-253.

Saltelli A, Chan K, Scott E., 2000. Sensitivity analysis. New York: John Wiley & Sons Inc.

Saltelli A., Tarantola, S., 2002. On the relative importante of input factos in mathematical models: safety assessment for nuclear waste disposal, Journal of the American Statistical Association, 97, 459, 702-709.

Schulz, T. L., 2006. Westinghouse AP1000 advanced passive plant, Nuclear Engineering and Design, vol. 236,pp 1547-1557.

Secchi, P., Zio, E., Di Maio, F., Quantifying Uncertainties in the Estimation of Safety Parameters by Using Bootstrapped Artificial Neural Networks, Annals of Nuclear Energy, Volume 35(12), Pages 2338-2350, doi: 10.1016/j.anucene.2008.07.010, 2008.

Yu, Y., Wang, S., Niu, F., 2013. Thermal–hydraulic performance analysis for AP1000 passive containment cooling system, Proceedings of the 21th International Conference on Nuclear Engineering ICONE21 July 29- August 2, 2013, Chengdu, Sichuan, China.

Yu, Y., Ma, G., Hao, Z., Wang, S., Niu, F., Zio, E. 2015, Correlation analysis for sceening key parameters for passive system reliability analysis, Annals of Nuclear Energy, 77, 23-29.

Wulf, W., Boyack, B.E., Catton, I., Duffey, R.B., Griffith, P., Kastma, K.R., Lellouche, G.S., Levy, S., Rohatgi, U.S., Wilson, G.E., Zuber, N., 1990. Quantifying reactor safety margins. Part 3: assessment and ranging of parameters. Nucl. Sci. Des.119, 33–65.

Zio, E. & Di Maio, F., 2008a. Bootstrap and Order Statistics for Quantifying Thermal-Hydraulic Code Uncertainties in the Estimation of Safety Margins, Science and Technology of Nuclear Installations, Article ID 340164, 9 pages, doi:10.1155/2008/340164.

Zio, E., Di Maio, F., Martorell, S., Nebot, Y., 2008b. Neural Networks and Order Statistics for Quantifying Nuclear Power Plants Safety Margins, Proceedings, European Safety & Reliability Conference (ESREL), Valencia, Spain.

Zio, E. & Pedroni, N., 2009. Estimation of the functional failure probability of a thermal-hydraulic passive system by Subset Simulation, Nuclear Engineering and Design, vol. 239, pp. 580-599.

Zio, E., Di Maio, F., Tong, J., 2010. Safety Margins Confidence Estimation for a Passive Residual Heat Removal System, Reliability Engineering and System Safety, RESS, Vol. 95, pp. 828–836, 2010.

APPENDIX A

	Input variable	Description	Units	Type of distribution	Mean value, μ	Standard Deviation, σ	μ-4σ	μ+4σ
1	G	Steam mass flow rate	kg/s	*normal*	182	5	162	202
4	T_{steam}	Steam temperature	°C	*normal*	163	8.15	130.4	195.6
5	t_{w1}	Containment wall thickness	m	*normal*	0.052	0.00026	0.051	0.053
6	t_{w2}	Containment wall thickness	m	*normal*	0.043	0.000215	0.04214	0.04386
7	t_{w3}	Containment wall thickness	m	*normal*	0.043	0.000215	0.04214	0.04386
8	t_{w4}	Containment wall thickness	m	*normal*	0.015	0.000075	0.0147	0.0153
9	t_{w5}	Containment wall thickness	m	*normal*	0.015	0.000075	0.0147	0.0153
10	D_1	Diameter of uphead	m	*normal*	43	0.215	42.14	43.86
11	H_1	Height of uphead	m	*normal*	13.51	0.06755	13.24	13.78
12	D_2	Diameter of containment	m	*normal*	43	0.215	42.14	43.86
13	H_2	height level of operative plant layer	m	*normal*	0	0.05	-0.2	0.2
14	H_3	Height level of the blade	m	*normal*	17.25	0.08625	16.905	17.595
15	H_4	Height level of the containment bottom	m	*normal*	-11	0.055	-11.22	-10.78
16	H_5	Containment height	m	*normal*	46.58	0.2329	45.65	47.51
17	D_3	Containment diameter	m	*normal*	43	0.215	42.14	43.86
18	V	Containment volume	m^3	*normal*	75000	750	72000	78000
19	A_1	Area in air baffle	m^2	*normal*	42.28	0.4228	40.5888	43.9712
20	A_2	Area in air baffle downcomer	m^2	*normal*	131.1	1.311	125.856	136.344
21	Z_1	Height of the downcomer air baffle	m	*normal*	32.51	0.16255	31.86	33.16
22	Z_2	Height of the riser air baffle	m	*normal*	32.51	0.16255	31.86	33.16
23	A_3	Area in air baffle riser	m^2	*normal*	125.8	1.258	120.768	130.832
24	A_4	Area in air baffle riser	m^2	*normal*	34	0.34	32.64	35.36
25	A_5	Area in air baffle riser	m^2	*normal*	156.97	1.5697	150.6912	163.2488
26	A_6	Area in air baffle downcomer intake	m^2	*normal*	113	1.13	108.48	117.52
27	A_7	Area at the blade turn	m^2	*normal*	68.49	0.6849	65.7504	71.2296
28	A_8	Area at diffusers level	m^2	*normal*	463.1	4.631	444.576	481.624
29	A_9	Area at the inlet of the chimney	m^2	*normal*	41.53	0.4153	39.8688	43.1912
30	A_{10}	Area at the exhaust of the chimney	m^2	*normal*	74.82	0.7482	71.8272	77.8128
42	H_c	Height of chimney	m	*normal*	8.27	0.04135	8.1046	8.4354
43	ρ_1	Central containment layer density	kg/m^3	*normal*	7750.476	38.75238	7595.47	7905.49
44	α_1	Central containment layer conductivity	W/(m K)	*normal*	51.9	5.19	31.14	72.66
45	Cp_1	Central containment layer heat capacity	J/(kg K)	*normal*	447.9876	2.239938	439.03	456.95
46	ρ_2	Covering layer density	kg/m^3	*normal*	3324.15	16.62075	3257.66	3390.63
47	α_2	Covering layer conductivity	W/(m K)	*normal*	0.52246	0.052246	0.31	0.73
48	Cp_2	Covering layer heat capacity	J/(kg K)	*normal*	544.284	2.72142	533.39	555.16

	Input variable	Description	Unit	Type of distribution	Lower value	Upper value
2	T_{inlet}	External air temperature	°C	*seasonal*	2	39
3	P_{inlet}	External air pressure	MPa	*seasonal*	0.0984	0.1011

	Input variable	Description	Unit	Type of distribution	Mean value, μ	Lower value	Upper value
31	f_1	Air baffle friction factor	-	*uniform*	1.15	1.035	1.265
32	f_2	Air baffle friction factor	-	*uniform*	3.74	3.366	4.114
33	f_3	Air baffle friction factor	-	*uniform*	1.6	1.44	1.76
34	f_4	Air baffle friction factor	-	*uniform*	0.5	0.45	0.55
35	f_5	Air baffle friction factor	-	*uniform*	1.13	1.017	1.243
36	f_6	Air baffle friction factor	-	*uniform*	0.5	0.45	0.55
37	f_7	Air baffle friction factor	-	*uniform*	3.9	3.51	4.29
38	f_8	Air baffle friction factor	-	*uniform*	1	0.9	1.1
39	f_9	Air baffle friction factor	-	*uniform*	3.68	3.312	4.048
40	f_{10}	Air baffle friction factor	-	*uniform*	2.76	2.484	3.036
41	f_{11}	Air baffle friction factor	-	*uniform*	1.27	1.143	1.397
49	T_{water}	Pool cooling water temperature	°C	*uniform*	25	1	32
50	u_{air}	Air baffle intake air speed	m/s	*uniform*	2	0.5	4
51	M_{vapor}	Initial in-containment vapor mass	kg	*uniform*	36600	33600	39600